Catalogue no. 12-593-XIE

Teacher's Guide to Data Discovery

Statistics Canada Statistique Canada

Canada

How to obtain more information

Specific inquiries about this product and related statistics or services should be directed to: Client Services, Culture, Tourism and the Centre for Education Statistics Division, Statistics Canada, Ottawa, Ontario, K1A 0T6 (telephone: 613-951-7608; toll-free at 1-800-307-3382; by fax at 613-951-9040; or e-mail: educationstats@statcan.ca).

For information on the wide range of data available from Statistics Canada, you can contact us by calling one of our toll-free numbers. You can also contact us by e-mail or by visiting our website at www.statcan.ca.

National inquiries line	1-800-263-1136
National telecommunications device for the hearing impaired	1-800-363-7629
Depository Services Program inquiries	1-800-700-1033
Fax line for Depository Services Program	1-800-889-9734
E-mail inquiries	infostats@statcan.ca
Website	www.statcan.ca

Information to access the product

This product, catalogue no. 12-593-XIE, is available for free in electronic format. To obtain a single issue, visit our website at www.statcan.ca and select Publications.

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, the Agency has developed standards of service which its employees observe in serving its clients. To obtain a copy of these service standards, please contact Statistics Canada toll free at 1-800-263-1136. The service standards are also published on www.statcan.ca under About us > Providing services to Canadians.

Acknowledgements

This guide was produced through Statistics Canada's Education Outreach Program under the guidance of Mary Townsend and edited by Elise Mennie. They gratefully acknowledge the work of Angela McCanny, an elementary teacher from Ottawa, Ontario, who undertook the initial writing of this guide to address the needs of her peers. She devoted considerable time and expertise to this project.

Special message to teachers

With the teaching community's help, Statistics Canada has created a vast array of curriculum-supportive activities, lessons and kits for most school subjects. These resources assist teachers in integrating relevant Canadian information into their classroom activities and in improving their students' statistical literacy. Visit the "Learning resources" section of the Statistics Canada website at www.statcan.ca.

2008-03304/

Statistics Canada
Culture, Tourism and the Centre for Education Statistics Division

Teacher's Guide to Data Discovery

Published by authority of the Minister responsible for Statistics Canada

© Minister of Industry, 2007

All rights reserved. The content of this electronic publication may be reproduced, in whole or in part, and by any means, without further permission from Statistics Canada, subject to the following conditions: that it be done solely for the purposes of private study, research, criticism, review or newspaper summary, and/or for non-commercial purposes; and that Statistics Canada be fully acknowledged as follows: Source (or "Adapted from", if appropriate): Statistics Canada, year of publication, name of product, catalogue number, volume and issue numbers, reference period and page(s). Otherwise, no part of this publication may be reproduced, stored in a retrieval system or transmitted in any form, by any means—electronic, mechanical or photocopy—or for any purposes without prior written permission of Licensing Services, Client Services Division, Statistics Canada, Ottawa, Ontario, Canada K1A 0T6.

June 2007

Catalogue no. 12-593-XIE
ISBN 978-0-662-44779-5

Frequency: occasional

Ottawa

La version française de cette publication est disponible sur demande (nº 12-593-XIF au catalogue).

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued cooperation and goodwill.

Table of contents

Introduction **5**
Essential skills for Canadian students 5
Statistical investigation 5

1.Choosing the dataset **6**
Primary data 6
What are primary data? 6
Resources for collecting primary data 6
Census at School 6
Statistics: Power from Data! 7
Households and the Environment Survey: School edition 7

Secondary data 7
What are secondary data? 7
Topics of interest that use secondary data 7
Sources of Canadian secondary data 8
E-STAT 8
Community Profiles 8
Summary tables 8
Health Indicators 8
Census data 8
The Daily 9
Publications 9
Microdata files 9
Sources of international secondary data 9
Census at School international data 9
E-STAT international data 9
International statistical websites 9

2. Understanding data concepts **10**
Predict and Investigate 10
Questions to ask students 10
Types of data 11
Categorical data 11
Numeric data 12
Types of graphs 13
Bar graphs 13
Line graphs 14
Circle graphs (pie charts) 15
Pictographs 16
Histograms 17
Scattergraphs 18
Measures of central tendency: Mean, median and mode 19

3. Analysing the data **20**
Data management tasks 20
Using computer software 20
Spreadsheet programs 21
Dynamic statistical software 21
Dynamic graphing software 22
Graphing calculator software 22
Where computers are not available 22
Activities that don't require computers 22

Introduction

This guide was written to support both elementary and secondary teachers in helping students develop basic statistical skills. It provides teachers with specific instructions on:

- finding interesting and grade-appropriate Canadian datasets
- choosing appropriate graphs for different kinds of data
- calculating basic statistical measures, with or without statistical software

Essential skills for Canadian students

Partners in Canadian education no longer question that we live in an information society. Teachers and parents recognize that students need to develop skills in a variety of technology-related areas in order to become well-educated citizens and contributing members of the work force. Indeed, skills in using computers, finding data and communicating with a high degree of numerical literacy are essential for Canadian students.

Teachers are finding that basic statistical skills – reading and creating tables, constructing graphs and calculating measures of central tendency (mean, median and mode) – are required across the curricula. Students practise these skills not only in the traditionally number-based subjects like mathematics and economics, but also in history, geography and other subjects. School boards across Canada are fostering these skills by allotting significant time for computer technology courses at the elementary level and offering new data management courses at the secondary level.

Many teachers already rely on Statistics Canada data to assist with these goals by accessing data in **E-STAT** and **Census at School**. They find that when students use Canadian data or data about themselves, they are drawn more deeply into thinking critically about the information. They compare their findings against their own experience; they question, draw conclusions and ultimately construct their own learning. With the growing availability of interesting data management software, students can quickly proceed to the analysis of the data, the step in which critical thinking takes place.

Statistical investigation

Teachers use data in the classroom for a variety of purposes. Sometimes, the goal is primarily to learn to work with data and draw conclusions from it. Sometimes, data are needed for exploring topics in the sciences, social sciences, or other subject areas. In both cases, two things are required to begin the process of statistical inquiry: a **question** to be answered and a **set of data**.

Whether students wish to first choose a question and then try to find data to support it or whether they work the other way around will depend upon the curriculum being explored.

The job of determining the question will be left to the teacher and students. However, to help with selecting suitable data, the following section presents various types of datasets available on the Statistics Canada website.

1. Choosing the dataset

Statistical investigation requires a set of suitable data. There are two categories of data: **primary data** and **secondary data**.

Primary data

What are primary data?

Primary data are collected to answer specific questions. They are gathered directly from first-hand sources by means of surveys, observation or experimentation.

Students may use **primary data** obtained through a ready-to-use survey such as Census at School (www.censusatschool.ca) or they can create their own survey or experiment to obtain data for a specific question. For example, the question, "Does the amount of time spent playing computer games affect school grades?" could lead students to design a classroom or school-wide survey on the connection between electronic gaming and study habits.

Resources for collecting primary data

The following resources available from Statistics Canada can help students collect primary data:

- **Census at School**

  www.censusatschool.ca

  Students complete an online survey in class, providing anonymous information about their lives and activities. The survey's questions were developed by an inter-provincial advisory group of teachers to be relevant to curricula and interesting to students. Some questions, such as languages spoken at home, fit the geography curriculum and resemble those on the national census. Other questions relate to the healthy living curriculum, such as smoking and nutrition or to social issues, such as bullying, role models and students' charities of choice.

  The following table of selected class results from the **Census at School** survey is a good example of primary data:

Table 1
Example of Census at School class results

Household	Language	Gender	Age	Height	Arm span	Eyes	Playing PC games	Reading	Watching TV	Pressure from school work
				centimetres			hours per week			
4	English	Male	13	172	173	Brown	20	7	9	some
3	English	Female	14	154	148	Brown	8	0	2	none
3	English	Female	13	166	161	Blue	0	6	11	some
3	English	Male	14	166	165	Brown	2	1	3	a lot
2	English	Female	14	157	150	Brown	4	0	2	a lot
6	English	Male	14	162	165	Other	15	0	5	some
4	English	Female	14	170	170	Green	7	4	14	some
5	English	Female	14	158	157	Blue	7	1	10	some

- **Statistics: Power from Data!**

  http://www.statcan.ca/english/edu/power/toc/contents.htm

  This guide to statistics will help students choose a specific question, plan a survey and analyse the collected data. Students will find practical advice about everything from questionnaire design and sampling methods to types of graphs and measures of central tendency. Included are exercises and answers for a vast number of statistical concepts, as well as lesson plans for teachers' use.

- **Household Environment Survey: School edition**

  http://www.statcan.ca/english/kits/houenv.htm

  Statistics Canada's Household Environment Survey measures the environmental practices and behaviours of Canadians. The school edition allows students to collect their own data using the same questionnaire. Students can compare their class results to those of the official Canadian survey.

Secondary data

What are secondary data?

Students may use data that have been gathered by someone else for another purpose but are suitable for their research needs. Such previously obtained information is called **secondary data**.

Secondary data may have been sorted, grouped into percentages, graphed or analysed in some way. They may be obtained from previously published materials, such as books, magazines, newspapers, government census publications and company reports and files.

Topics of interest that use secondary data

Statistics Canada provides free online access to a wide variety of Canadian secondary data. Teachers will find data tables that support research about a wide variety of topics, including some that may be of specific interest to young people. For example, the Health Indicators tables (listed below under Sources of Canadian secondary data) provide information about youth and health issues such as:

- age-specific mortality rates
- alcohol consumption
- food consumption
- physical activity
- teenage pregnancies
- tobacco use

Students have used Statistics Canada resources to create interesting projects on various topics, such as those whose titles are shown below:

- Trends in Energy Consumption in Canada
- Teen Pregnancy and Abortion
- Obesity and Diabetes: A Growing Epidemic
- Teenager Driving Infractions
- Education, Salary and Career Paths

Sources of Canadian secondary data

The following resources may be used to find topics related to Canadian curricula and student interests.

- **E-STAT**

  http://estat.statcan.ca

  For math and social sciences, use E-STAT to graph, tabulate, download, analyse and map a huge selection of data. It is Statistics Canada's most comprehensive source of data and covers the widest span of years and most detailed levels of geography. Developed specifically for schools, E-STAT contains chronological data from hundreds of social and economic surveys, as well as several hundred census variables for every province, county and municipality in Canada, right down to census tracts or neighbourhoods. Of particular interest to students are data on the health behaviour of school-age children. Free school registration is required. User name and password are required for use outside the school.

- **Community Profiles**

  http://www12.statcan.ca/english/census06/data/profiles/community/Index.cfm?Lang=E

  Visit your community! As data from the 2006 Census become available, they are added to the *2006 Community Profiles* for every community in Canada. Five major topics are included: population; families and dwellings; education; work; earnings and income. Find 2001 data on all these topics by consulting the *2001 Community Profiles*, which contain over 150 variables from the Census of Population, Census of Agriculture and Aboriginal Peoples Survey.

- **Summary tables**

  http://www40.statcan.ca/l01/cst01/

  Summary tables, formerly called Canadian statistics, provide continuously updated national and provincial data gathered through hundreds of surveys. These top 500 tables are sorted into 32 broad subject categories. You can search the tables by subject, province, territory or metropolitan area. The Overview section offers useful brief articles related to each subject.

- **Health Indicators**

  http://www.statcan.ca/bsolc/english/bsolc?catno=82-221-X

  Over 80 indicators measure the health of the Canadian population and the effectiveness of the health care system. Designed to provide comparable information at the health region and provincial/territorial levels, these data are produced from a wide range of the most recently available sources.

- **Census data**

  http://www12.statcan.ca/english/census/index.cfm

  While E-STAT and Community Profiles are the primary sources of census data for students, the Census section of Statistics Canada's website provides additional resources, including information about the census, articles from the Analysis series, data tables organized by topic and geography, maps, animated presentations and reference documents.

- **The Daily**

  http://www.statcan.ca/english/dai-quo/

  The Daily news releases cover a wide range of topics and usually include tables and/or graphs. Use The Daily's search engine to find articles of interest and then sort using date to find the most recent ones. Articles often refer to CANSIM table numbers that allow you to further investigate the data. Find these tables free in E-STAT. Articles also link to a section called 'Definitions, data sources and methods: survey number', where you can read about sampling techniques and view the actual questionnaire.

- **Publications**

  http://www.statcan.ca/english/pub/index.htm

  This area provides lists of publications: electronic versions are available for free and print versions are for sale. Subject-specific data tables can be found within publications on the environment, health, education, justice and other topics.

- **Microdata files**

  http://www.statcan.ca/english/kits/courses/math2.htm

  Microdata files are a good source of large datasets suitable for manipulation using dynamic statistical software (such as Fathom). They provide a representative sample of anonymous individual survey responses. For examples, see the following microdata sources:
  - 2001 Census
  - National Longitudinal Survey of Children and Youth

Sources of international secondary data

- **Census at School international data**

  http://www.censusatschool.org/international

  In addition to providing primary data about your own class, Census at School allows students to access data collected from a number of countries, including the United Kingdom, Canada, New Zealand, Australia and South Africa. These are useful for questions comparing international conditions.

- **E-STAT international data**

  http://estat.statcan.ca/content/english/hbsc/hbsc_start.shtml

  Although most of its data are Canadian, E-STAT also contains some information from other countries. Of particular interest to students are the data on the health behaviours of school-age children from different countries.

- **International statistical websites**

  http://www.statcan.ca/english/reference/othsit.htm

  Students can use these links to national statistics offices and other international statistical organizations to search trends in dozens of countries for world issues projects.

2. Understanding data concepts

Predict and investigate

As discussed in the introduction, a statistical investigation sometimes starts with a student asking a question that leads to the gathering of the necessary primary data. At other times, a secondary set of data is available that can lead the student to pose interesting questions.

In either case, students make better sense of the data if they are led to think about possible answers before, during and after the data collection. Students must learn to **predict and investigate**.

Once a question has been posed, students must begin to think about how they can find the answer to it. A good starting place is to ask students to **predict** or **hypothesize** what they think the answer will be, and then see how that leads to the **investigation** of the problem.

Let's consider the following research question: Does the amount of time spent playing computer games affect school grades?

Students may initially respond 'no' or 'yes' to this question. By asking them to explain their answers, you can encourage them to think about the conditions they would attach to their answer.

For example, they may suggest that the answer depends on the age or gender of students. Perhaps they'll think the number of hours spent on homework has an effect, as well as the time spent playing computer games. These ideas, in turn, may suggest the need to include questions about homework as well as computer games in their survey. Examining their initial predictions will lead them to the steps of investigation.

Additionally, when you ask students to predict an answer before they investigate the question, you help them to notice and correct any misconceptions as they are collecting their data.

Questions to ask students

- What do you think the answer to your question will be?
- Are there any exceptions to your prediction?
- What data will you gather?
- How will you gather the data?
- Who will you ask?
- What measurement system will you use?
- How will you record the information?

Types of data

Particular questions produce particular types of data, which in turn lend themselves to particular types of graphs.

There are two main types of data: **categorical** and **numeric**.

Categorical data

The question 'What colour is your hair?' produces **categorical** data, which fit into the categories 'brown,' 'blonde,' 'black,' 'red' or 'other.' Categorical data can be broken down into **nominal** and **ordinal** sub-types.

See the table below for each categorical sub-type and its associated graph types.

Table 2
Categorical data

Types of data	Sub-types	Examples from Census at School database	Appropriate graphs
Categorical: Data fit into various categories of responses to a question.	**Nominal:** These data are identified by particular names or categories. These data cannot be organized according to any 'natural' order.	**Gender:** male, female **Favourite subject:** math, history, gym, music, etc. **Eye colour:** brown, blue, green, other **Pets:** cats, dogs, birds, fish, etc.	Bar graph, circle graph, pictograph
	Ordinal: These data are identified by categories that can be placed in a specific order or ordered in some 'natural way.'	**Schoolwork pressure:** none, very little, some, a lot	Bar graph, circle graph, pictograph

Numeric data

The question, 'How many people live in your home?' produces numeric data, which can be broken down into discrete and continuous sub-types.

See the chart below for each sub-type and its associated graph types.

Table 3
Numeric data

Types of data	Sub-types	Examples from Census at School database	Appropriate graphs
Numeric: Data are represented by real numbers. Also known as quantitative data.	**Discrete:** Data that can only assume a finite number of different responses. For example, the numbers of people in a household are discrete data because you can only answer using whole numbers from 1 to 10 or more. You cannot include all the decimals or fractions in between as possible answers. For example, it's impossible to have 2.5 or 3.75 people.	**Age in years:** 7, 8, 9, 10, 11, etc. **Number of people in the household:** 1, 2, 3, 4, 5, etc. **Number of days during which you did an intense physical activity last week:** 0, 1, 2, 3, 4, 5, 6, 7, etc.	Bar graph, line graph, circle graph, histogram
Note: Sometimes numbers can represent scales of response (e.g., 0=none, 1=very little, 2=some, etc.). In this case, the responses are considered ordinal categorical data, not numeric data, even though they are represented by a number.	**Continuous:** Data that can assume an infinite number of different responses. The answers have infinite possibilities since they can include decimal responses. For example, a student's height may be 1.57923 metres.	**Height, arm span, wrist circumference:** It's impossible to list all the possibilities. Note: In the Census at School survey, students are required to round their answers to the nearest centimetre or millimetre, so in effect their responses are discrete data.	Line graph, histogram

Notes: To make **continuous data** easier to handle, they are often grouped into class intervals. Grouping data is part of the process of organizing data so that the information becomes useful. For example, instead of displaying every height measured in a class of students, it is more effective to display grouped categories such as 120 to 129 cm, 130 to139 cm, 140 to 149 cm, etc.

Discrete data may be grouped or ungrouped. Grouping data makes them easier to handle, but with a small number of responses, it can be just as clear to leave them ungrouped.

Types of graphs[1]

Bar graphs

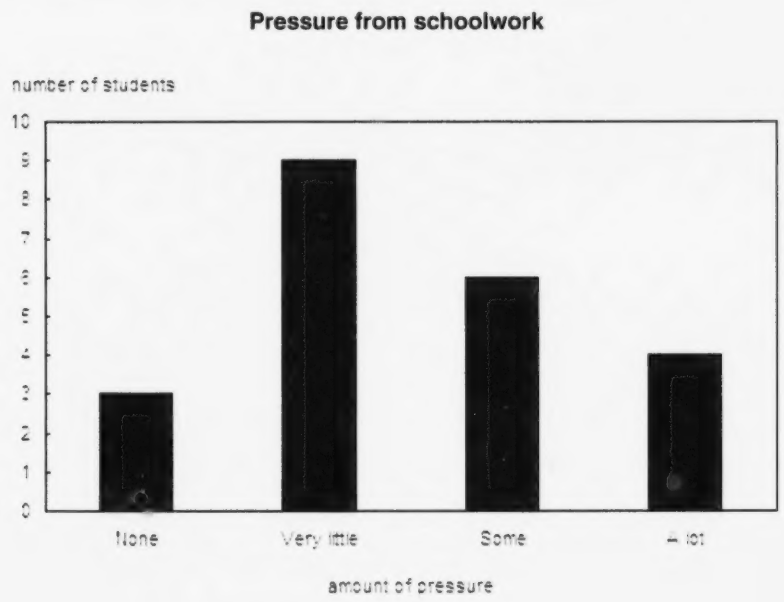
Bar graphs can present either **categorical** or **numeric** data. Numeric data are either **ungrouped** (if they include few numbers) or **grouped** into class intervals.

Bar graphs consist of an axis with labelled horizontal or vertical bars. Those with vertical bars are also called column graphs. The bars depict the frequencies of different responses. The numbers on the x-axis of a horizontal bar graph or the y-axis of a vertical bar graph are called the **scale**.

When developing bar graphs, each category or value is represented by a vertical or horizontal bar. The height or length of the bar will represent the number of units or observations in that category (i.e. their frequency).

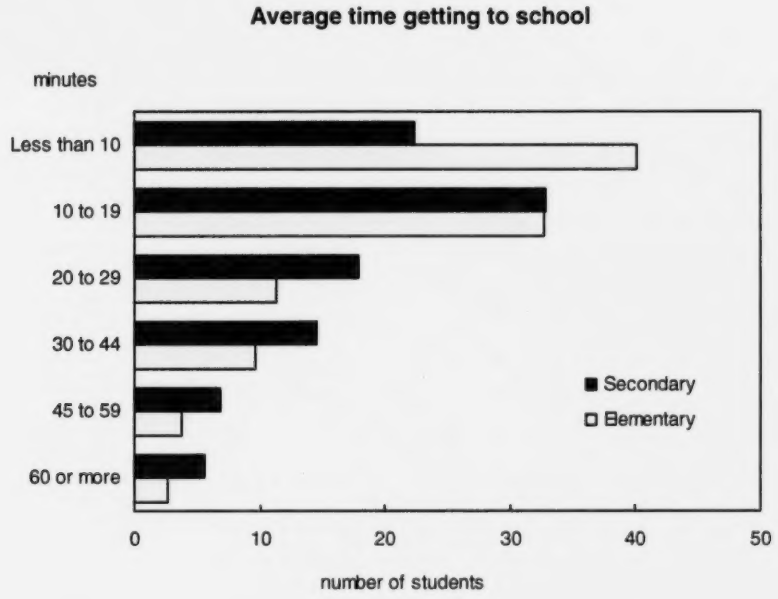
Three-dimensional bar graphs should be avoided because the added depth dimension makes it more difficult to read the data accurately.

Figure 1 Example of vertical bar graph:



1. For more information, see *Statistics: Power from Data!*, Graph types, at http://www.statcan.ca/english/edu/power/ch9/first9.htm

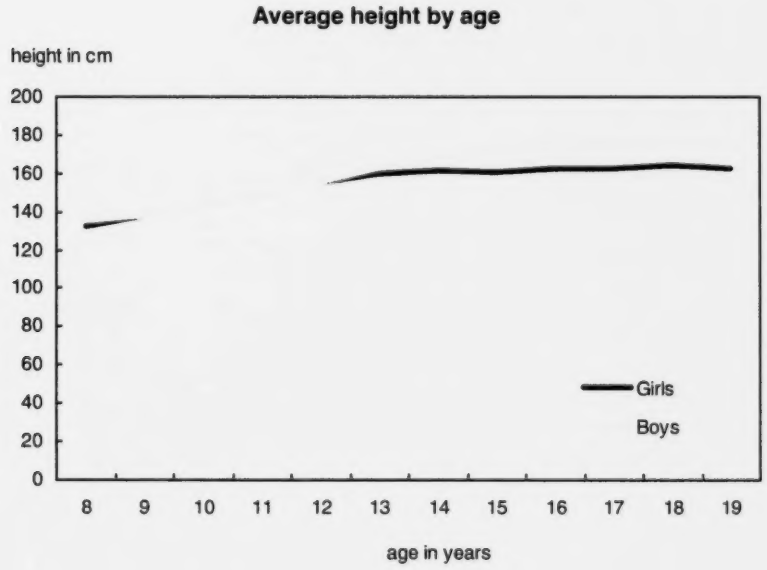
Figure 2 Example of horizontal bar graph:



Line graphs

Line graphs compare two variables: one is plotted along the x-axis (horizontal) and the other along the y-axis (vertical). The graph shows how the variables are related or vary with each other by drawing a continuous line between all the points.

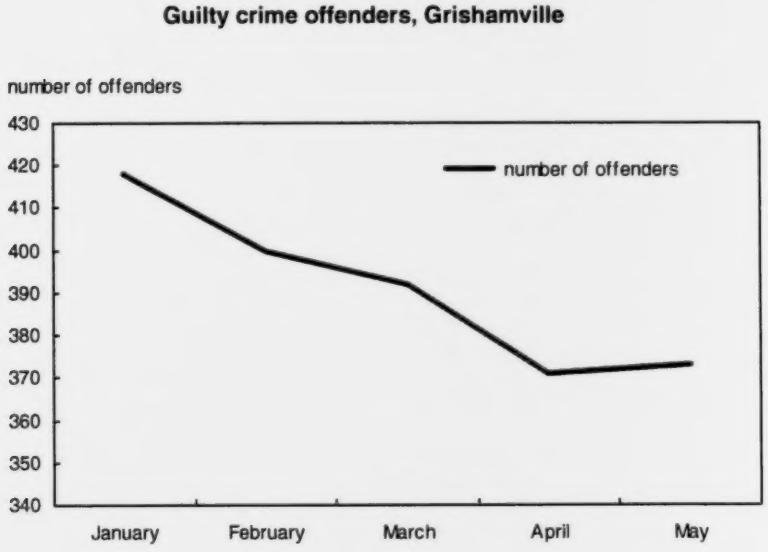
Figure 3 Example of line graph showing relationship between two variables:



Line graphs are also used to reveal trends over time. While bar graphs reveal a change in magnitude, line graphs show a change in direction. Line graphs are popular for showing data over time because they reveal data trends clearly and are easy to create.

When a line graph is showing a trend over time, the y-axis usually indicates quantity (e.g., dollars, litres) or percentage, while the horizontal x-axis measures units of time.

Figure 4 Example of line graph showing trend over time:

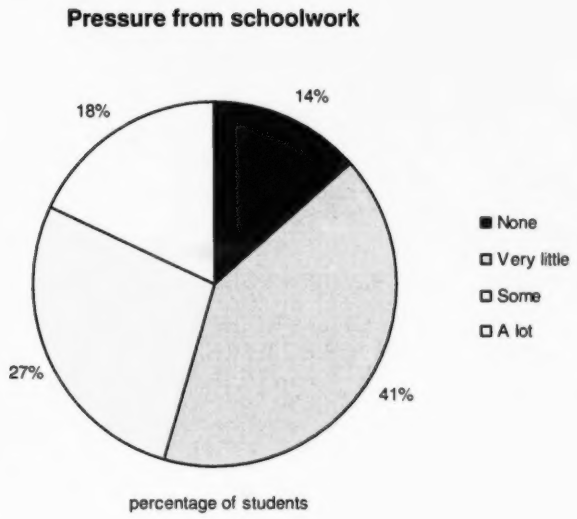


Circle graphs (pie charts)

A circle graph or pie chart is a way of summarizing a set of **categorical data** or displaying the different values of a given variable (e.g., percentage distribution). This type of graph is a circle divided into segments, with each segment representing a particular category and its proportion of the total. The area of each segment is the same proportion of a circle as the category is of the total data set.

Circle graphs are best used when there are few categories – ideally no more than six – otherwise, the resulting picture will be too complex to understand. Never use a three-dimensional pie chart, even when it's available as a graph option in spreadsheet software. The 3-D image is misleading because the surface area of some segments can appear larger than the actual proportions they represent.

Figure 5 Example of circle graph:



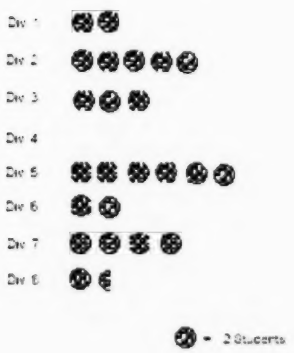
Pictographs

A pictograph uses picture symbols to convey the meaning of **categorical data**. It is similar to a bar graph in that each horizontal or vertical row represents the frequency or number of responses in each category. Pictographs should be used carefully because the pictures may, either accidentally or deliberately, misrepresent the data.

For example, the cookie image in the pictograph below represents two students and the half-cookie image represents one student. Other types of pictographs may use an image that grows larger or smaller to represent changes in data. In such cases, care must be taken to ensure the size or area (total surface) of the picture is proportional to the change it is representing.

Figure 6 Example of pictograph:

Number of students who like chocolate chip cookies best



Histograms

A histogram is used to summarize either **discrete** or **continuous numerical data** that are measured on an interval scale. It is often used to illustrate major features of the distribution of the data. A histogram divides the range of possible values into classes or groups. For each group, a rectangle is constructed with a base length equal to the range of values in that specific group and an area proportional to the number of observations or frequency of that group. This means that the rectangles will be drawn of non-uniform height. A histogram has an appearance similar to a vertical bar graph, but when the variables are continuous, there are no gaps between the bars. When the variables are discrete, however, gaps should be left between the bars.

Figure 7 Example of histogram illustrating continuous data:

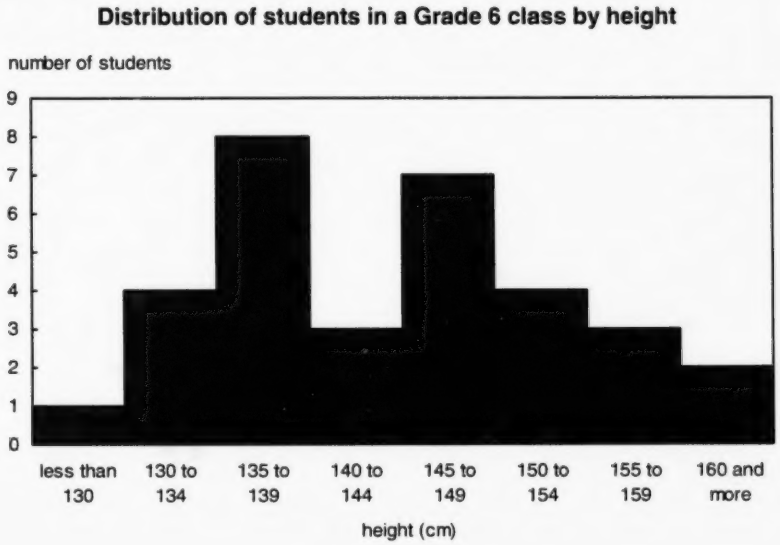
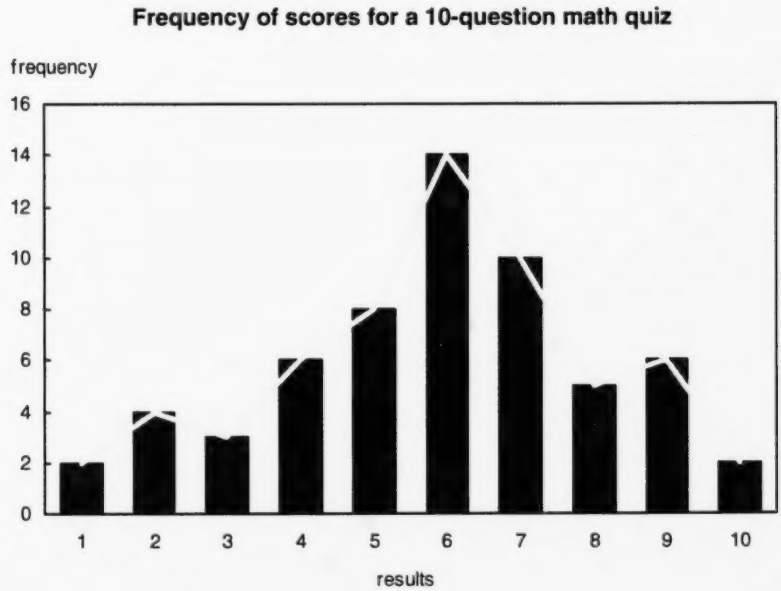


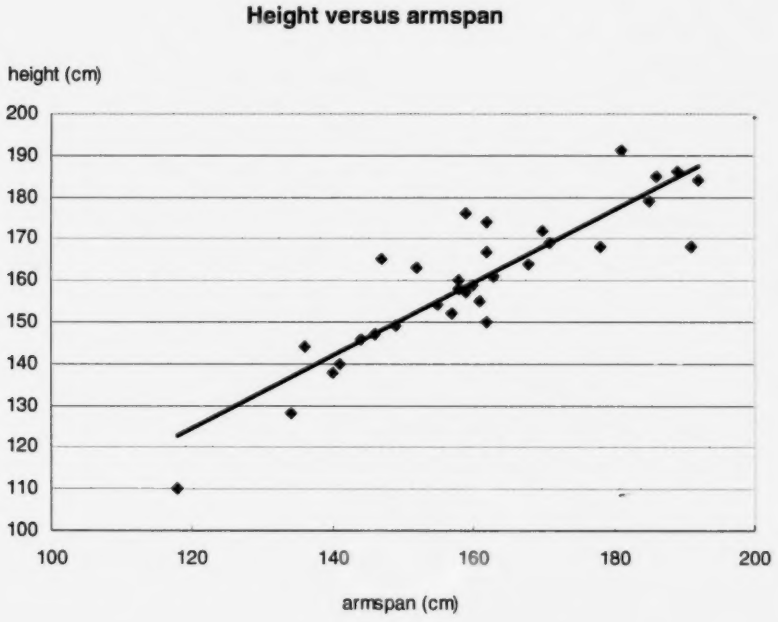
Figure 8 Example of histogram illustrating discrete data:



Scatter graphs

Scatter graphs are used to show a relationship between two variables by means of ordered pairs plotted on a coordinate grid. The data points are not joined; the resulting pattern indicates the type and strength of the relationship between the variables. A line of best fit can be drawn between the points when a relationship exists. Scatter graphs can illustrate data correlation, positive or negative relationships between variables, non-linear patterns, spread of data and outliers.

Figure 9
Example of a scatter graph showing a positive correlation:



Measures of central tendency: Mean, median and mode[2]

The mean, median and mode can help you capture, with a single number, what is typical of a set of data. For example, a typical Grade 8 class can be composed of 12 to 15 year olds. However, if we find more 13 year olds than any other age group, we use the **modal** age 13 to represent the age of grade eight students in that particular class. Depending on the situation, either the mean, the median or the mode may give the best description of a particular set of data.

The **mean** is the **average** value in a data set. It is calculated by adding all the data and dividing the sum by the total number of data items in the set.

The **median** is the **middle** value in a data set that has been arranged in numerical order – exactly half the data are above the median and half are below it. You must first arrange the data in ascending or descending order to determine its middle number. If there is an even number of data, you must average the two middle numbers to find the median.

The **mode** is the value that occurs **most frequently** in the set. When two numbers occur equally frequently the data are bi-modal.

In a **normal distribution**, the mean, median and mode are identical in value. For example, the following data set shows a normal distribution:

Data set:	14, 14, 13, 15, 15, 14, 13, 14, 13, 15
mean:	(14 + 14 + 13 + 15 + 15 + 14 + 13 + 14 + 13 +15) / 9 = 14
median:	the median is 14 (13, 13, 13, 14, **14**, **14**, 14, 15, 15, 15)
mode:	the most frequent number is 14

2. For more information, see *Statistics: Power from Data!*, Measures of central tendency, at http://www.statcan.ca/english/edu/power/ch11/first11.htm

3. Analysing the data

Data management tasks

Once a set of data has been obtained, it is time to begin analysing it. This means engaging in some or all of the following activities:

- organizing the data into tables
- calculating percentages
- calculating measures of central tendency
- constructing graphs

In addition to performing these mechanical tasks, students are required to think about the meaning of the data.

For example, they can explore the relationship between two variables or observe the effect that removing extreme values has upon the mean or the median. Most provincial curricula for Grades 7 and 8 specify that students will engage in a set of skills in the data management strand which include

- collecting and organizing categorical, discrete and continuous data
- displaying data in relative frequency tables, bar graphs, circle graphs, histograms, etc.
- identifying bias in data
- relating changes in data to changes in central tendency
- making inferences based on data

Using computer software

Many software programs will do most of the mechanical work of creating tables, performing calculations and constructing graphs with just a few clicks of the mouse. By eliminating the slow and often tedious work that can act as a barrier, these programs allow students to quickly progress to deriving meaning from the data.

Students can use computer software to question the data and experiment with different types of graphs that represent them or to examine the relationships between variables. This means that within a single class period, students as young as Grade 4 can find meaning in the data without having to spend the entire time working on a single graph. Additionally, since most students enjoy using technology, they are having fun while they are working.

If you have limited access to technology, some of the best exploration of data comes in small group settings where students can ask each other their 'I wonder if ...' questions and work through them as a community of learners. For this reason, you don't need to have one computer per student. Many data management activities lend themselves to learning in pairs or groups of four. This means that four to six computers can be enough for a lesson.

Many software programs are used by the various school boards across Canada. The list below provides an overview of the most commonly used programs, but is by no means complete. Please note that **Statistics Canada** does not endorse the use of any particular program.

Software for data management can be categorized into these subgroups:

- spreadsheet programs
- dynamic statistical software
- dynamic graphing software
- graphing calculator software

Spreadsheet programs

Examples:

- Excel (Grades 4 to 12)
  http://office.microsoft.com/en-us/FX010858001033.aspx
- QuattroPro (Grades 4 to 12)
  http://www.functionx.com/quattropro/lesson01.htm
- AppleWorks (Grades 4 to 12)
  http://www.apple.com/AppleWorks/tips/_nak.ssbasic/
- Calc in Open Office (Grades 4 to 12)
  http://www.tutorialsforopenoffice.org/category_index/spreadsheet.html

Data imported into a spreadsheet program are displayed in a table consisting of rows and columns. Quick calculations of data ranges, column totals, means and percentages can be obtained by entering formulae into the appropriate cells. Some programs allow one-click options for certain calculations.

To graph the data, use the graphing tool provided in your spreadsheet program. For categorical data, you must first make a frequency table before creating the graph. In addition, Excel also offers a **pivot chart** option.

The graphing tool lets students select a type of graph and personalize it with their choice of colour, font, type size, etc. See **Types of Data** on Page 8 of this document to find appropriate types of graphs for different types of data.

Dynamic statistical software

Examples:

- TinkerPlots (Grades 4 to 8)
  http://www.keypress.com/x5715.xml
- Fathom (Grades 8 to 12)
  http://www.keypress.com/x5656.xml

Data are imported into these programs by cutting and pasting a URL address or by importing a spreadsheet file in comma separated values (.csv) or plain text (.txt) formats. The data appear as a list of attributes in a data card, which you can transfer into a graph by the drag-and-drop method. Each point on the graph represents the response of one individual in the sample. That individual's entire set of attributes can be brought up by clicking on the data point.

Students may pose a question and then create one or two different graphs simply by dragging columns of data onto the axes. They can pose a different question, then drag in a different column and see the changes immediately.

These are some of the graph types available in these programs:

- Dot plot (stacked or unstacked)
- Line plot
- Histogram
- Box plot
- Bar chart
- Scatter plot

Dynamic graphing software

Example:

- EasyGraph (Grades 8 to 12)
  http://www.dahwin.com/newpage1.htm

Graphing software is useful for exploring relationships between two variables. Data are entered point by point on a Cartesian grid (x- and y-axes), creating a scatter plot diagram. Program features include finding the line of best fit through the points and obtaining an instant summary of statistical measures (mean, variance and standard deviation). Students can investigate the effect that eliminating outliers may have on the shape of the graph or the mean.

Graphing calculator software

Example:

- TI-Interactive or TI-Connect for use with TI graphing calculator (Grades 7 to 12)
  http://education.ti.com/educationportal/sites/US/homePage/index.html

Graphing calculators allow students to explore the relationship between two variables. Students can either manually enter the data for each variable into a table of values or import them using a link cable.

The graphing calculator can graph data as a scatter plot and will provide lines of best fit and statistical measures (mean, variance and standard deviation). Most provincial curricula require the use of graphing calculators for certain math courses from Grades 8 to 11.

Where computers are not available

In many schools, computers and software programs are not always available, so it is necessary to do much of the work by hand. Even when computers are readily available, it may be preferable for students to do the work by hand for certain parts of the curriculum. By going through the manual steps themselves, students can understand the concepts involved in creating tables and graphs and calculating statistical measures.

Many valuable activities are available for working with data by hand. The list below offers a selection of them that you can print or download from the Internet. Note that in some cases, you will also need to print selected data from a spreadsheet before beginning the activity so that the students can have a hard copy to work with.

Activities that don't require computers

- **You are the researcher!** (Grades 4 to 8)
  http://www19.statcan.ca/02/02_037_e.htm
  This mini-project asks students to decide on an interesting question to research using their Census at School class results. They then display the results using different types of graphs.

- **Worksheets for analysing class data** (Grades 9 to 12)
  http://www19.statcan.ca/02/02_046_e.htm
  These activities use Census at School class data to cover frequency tables, measures of central tendency, circle graphs, histograms, lines of best fit, etc. A marking rubric is included.
  Note: In the last activity, students are asked to compare their handmade line of best fit with one created on a graphing calculator. You can easily omit this section if graphing calculators are not available.

- **Other Census at School learning activities** (Grades 4 to 12)
  http://www19.statcan.ca/02/02_004_e.htm
  These activities make use of Census at School class results in the math, social science and health curricula. While some lessons ask students to retrieve Canadian or international data from the website using computers, most only require that the teacher print selected columns from the class data prior to the activity.

  Examples of lesson titles:

  Travel to school (Grades 4 to 8): Analyse the data using stem and leaf plots and pie charts.

  What a zoo! (Grades 4 to 8): Examine bar graphs that represent the same data but use different scales.

  The Vitruvian theory—does it apply to you? (Grades 4 to 8) and (Grades 9 to 12): Verify the theory described by Leonardo Da Vinci that arm span is equal to height.

- **E-STAT lesson plans** (Grades 7 to 12)
  http://www.statcan.ca/english/Estat/lesson.htm

  These activities make use of E-STAT data in the math, social science and health curricula. Complete student instructions and worksheets are included. While most lessons in this section ask students to use computers to retrieve data tables from E-STAT, the teacher can easily print the tables and distribute them to students before the activity.

  Examples of lesson titles:

  Comparing the food choices and body image of 15-year-olds around the world (Grades 7 to 12): Examine and compare data from various countries.

  Role playing Jean Talon (Grades 7 to 9): Taking the role of Jean Talon, use 1665 census data to convince the King of France to increase investments in New France.

  Smoking today = smoking tomorrow (Grades 7 to 9): Research trends by examining smoking data by age and gender.